--1 (once amended). A method of detecting a facial region within a video comprising the steps of:

(a) receiving a first frame of said video comprising a plurality of pixels;

(b) receiving a subsequent frame of said video comprising a plurality of pixels;

(c) calculating a difference image representative of the difference between a plurality of said pixels of said first frame and a plurality of said pixels of said subsequent frame;

(d) determining a plurality of candidate facial regions within said difference image based on a transform of said difference image in a [spacial] spatial domain to a parameter space; and

(e) fitting said plurality of candidate facial regions to said difference image, where said difference image used for said fitting is free from being transformed as a result of step (d), to select one of said candidate facial regions.

--10 (once amended). A method of detecting a facial region within a video comprising the steps of:

(a) receiving a first frame of said video comprising a plurality of pixels;

(b) receiving a subsequent frame of said video comprising a plurality of pixels;

(c) calculating a difference frame representative of the difference between a plurality of said pixels of said first frame and a plurality of said pixels of said subsequent frame;

(d) determining a plurality of candidate facial regions within said difference frame; and

(e) fitting said candidate facial regions to said difference image to select one of said candidate facial regions based on a combination of at least two of

the following three factors including, a fit factor representative of the fit of said candidate facial regions to said difference image, a location factor representative of the location of said candidate facial regions within said [video] difference image, and a size factor representative of the size of said candidate facial regions.

--15 (once amended). A method of determining sensitivity information for a video comprising the steps of:

(a) receiving a first frame of said video;

(b) receiving a subsequent frame of said video;

(c) determining a spatial location of a facial region within said video based on at least said first and said subsequent [frames] frame; and

(d) calculating a sensitivity value for each of a plurality of spatial locations within said video based upon both said spatial location of said facial region within said video in relation to said spatial plurality of locations and a non-linear model of the sensitivity of a human visual [system] system's ability to perceive image detail at eccentric visual angles.

--18 (Amended Once). The method of claim 16 wherein said eccentricity is derived according to the following,

$$\theta_E = \frac{180}{\pi} \tan^{-1}\left(\frac{\sqrt{\left(\frac{y - y_C}{y_R}\right)^2 + \left(\frac{x - x_C}{x_R}\right)^2} - 1}{V} \right)$$

where θ_E is said eccentricity, y is a vertical pixel position within said video, x is a horizontal position within said video, x_c represents a horizontal component of a center position of an elliptical said facial region, y_c represents a vertical component of said center position of said elliptical said facial region, x_r represents a first elliptical radii of said elliptical said facial

feature in a horizontal direction; y_r represents a second elliptical radii of said elliptical said facial feature in a vertical direction, and V represents a viewing distance.

--20 (Amended Once). The method of claim 17 wherein said sensitivity versus eccentricity non-linear model is derived according to the following,

$$S = \frac{1}{1 + k_{ECC}\theta_E}$$

where S is representative of said sensitivity, k_{ECC} is a constant, and θ_E is representative of a non-linear contrast sensitivity function.

--21 (once amended). A method of encoding a video comprising the steps of:

(a) receiving a frame of said video consisting of a plurality of pixels;

(b) calculating sensitivity information for a plurality of locations within said frame of said video calculated based upon the sensitivity of a human visual system of a viewer [observing] perceiving image detail at eccentric visual angles of a particular region of said frame of said video, where said particular region of said frame is determined based upon the content of the frame itself; and

(c) encoding said frame in a manner that provides a substantially uniform apparent quality to perceiving detail at eccentric visual angles of said plurality of locations of said frame to said viewer when said viewer is observing said particular region of said frame of said video.

--22 (once amended). The method of claim 21 wherein said encoding of each of said plurality of locations of said frame of said video is based on a respective quantization value representative of a base quantization factor divided by said sensitivity information for a respective one of said plurality of locations in a manner that said encoding employs at least two

different quantization values, where said plurality of locations within said video are determined based upon the content of the frame itself.

--23 (once amended). The method of claim 22 wherein said encoding is derived in accordance with the following:

$$Q/S_1, Q[,]/S_2, Q/S_3,..,Q/S_N$$

where Q is representative of said base quantization factor, and S_1 through S_N are representative of said sensitivity information for said plurality of locations.

--26 (once amended). The method of claim [24 wherein S_k is a maximum of said plurality of said sensitivity information] 21 wherein said encoding of said frame of said video includes at least two different quantization values.

--30 (Amended Once). The method of claim 29 wherein said base quantization factor is derived in accordance with the following:

$$Q' = \sqrt{\frac{-AK}{B - ANC}\sum_{i=1}^{N}\sigma_i^2 S_i^2}$$

where A is representative of the number of pixels in one of said plurality of locations, K and C are constants associated with said plurality of locations, N is representative of the number of said plurality of locations, B is representative of said total number of bits, the σ_i^2 values are a measure how much texture is associated with said plurality of locations, and the S_i^2 values are representative of the respective said sensitivity information squared.

--31 (once amended). A method for encoding multiple blocks in a frame of image data, comprising:

(a) identifying a target bit value equal to a total number of bits available for encoding the frame;

(b) calculating sensitivity information for each one of the blocks based upon the sensitivity of a human visual system [observing] perceiving image detail at eccentric visual angles of a particular region of the image, where said eccentricity of said particular region of said image is determined based upon the content of the frame itself;

(c) adapting quantization values for each of the multiple blocks to provide substantially uniform apparent quality to perceiving detail at eccentric visual angles of each of the blocks in the frame subject to a constraint that the total number of bits available for encoding the frame is equal to the target bit value; and

(d) encoding the blocks with the quantization values.

A7 CON'T

--32 (Amended Once). The method of claim 31 wherein the quantization values are derived according to the following,

$$Q' = \sqrt{\frac{AK}{B - ANC}\sum_{i=1}^{N}\sigma_i^2 S_i^2}$$

where, Q_i is the quantization value for each block i, N is the number of blocks in the frame, B is the total number of bits available for encoding the frame, A is a number of pixels in each of the multiple blocks, K and C are constants associated with the image blocks, σ_i is an empirical standard deviation of pixel values in the block, and S_i is a weighting incorporating the sensitivity information for the block.

A8

--37 (once amended). A method for encoding video comprising the steps of:

(a) detecting the location of a facial region of a frame of said video;

(b) calculating a sensitivity value for each of a plurality of locations within said frame of said video based upon said location of said facial region; and

(c) encoding said frame in manner that provides a substantially uniform apparent quality to perceiving detail at eccentric visual angles of said plurality of locations to said viewer when said viewer is observing said facial region of said video.

A8
CON'T

--38 (once amended). The method of claim 37 wherein said sensitivity values are calculated based upon a non-temporal said location of said facial region, a non-temporal size of said facial region, and a non-linear model of the [sensitivity of a] human visual [system] system's ability to perceive image detail at eccentric visual angles.

--49 (Amended Once). The method of claim 48 wherein said eccentricity is derived according to the following,

A9

$$\theta_E = \frac{180}{\pi} \tan^{-1}\left(\frac{\sqrt{\left(\frac{y - y_C}{y_R}\right)^2 + \left(\frac{x - x_C}{x_R}\right)^2} - 1}{V} \right)$$

where θ_E is said eccentricity, y is a vertical pixel position within said video, x is a horizontal position within said video, x_c represents a horizontal component of a center position of an elliptical said facial region, y_c represents a vertical component of said center position of said elliptical said facial region, x_r represents a first elliptical radii of said elliptical said facial

AA
Cont'd

feature in a horizontal direction; y_r represents a second elliptical radii of said elliptical said facial feature in a vertical direction, and V represents a viewing distance.

--54 (once amended). The method of claim 53 wherein said encoding is derived in accordance with the following:

AID

$$Q/S_1,\ Q[,]/S_2,\ Q/S_3,..,Q/S_N$$

where Q is representative of said base quantization factor, and S_1 through S_N are representative of said sensitivity information for said plurality of locations.--

--61 (Amended Once). The method of claim 60 wherein said base quantization factor is derived in accordance with the following:

All

$$Q' = \sqrt{\frac{AK}{B - ANC}\sum_{i=1}^{N}\sigma_i^2 S_i^2}$$

where A is representative of the number of pixels in one of said plurality of locations, K and C are constants associated with said plurality of locations, N is representative of the number of said plurality of locations, B is representative of said total number of bits, the σ_i^2 values are a measure how much texture is associated with said plurality of locations, and the S_i^2 values are representative of the respective said sensitivity information squared.--

REMARKS

A replacement specification is submitted herewith to conform the pending specification to the specification in the parent application, serial number 09/052,591 filed 3/31/1998. The replacement specification merely conforms the text to the parent application, and in particular includes the formulas which were inadvertently omitted. No new matter has been added.